DOCUMENT RESUME

ED 137 782 CS 203 271

AUTHOR Bay, Libby; McCulloch, Elizabeth
TITLE Towards Uniformity in Grading Standards.
PUB DATE 76
NOTE 8p.; Paper presented at the Annual Meeting of the New York State English Council (26th, October 15-17, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Academic Standards; College Freshmen; Community Colleges; *Composition (Literary); *English Departments; *English Instruction; Essay Tests; *Expository Writing; *Grading; Junior Colleges

ABSTRACT

To study grading standards and consistency within the English department, 1600 freshmen at Rockland Community College were asked to complete a uniform exit essay at the end of English 101. After developing criteria for grading the papers, members of the department marked their own papers and one other set. Eight months later, 240 of the papers were regraded by the original instructor, in order to assess self-consistency in marking. Comparison of final grades, essay grades assigned by the instructor, and essay grades assigned by the disinterested marker suggested that there was a general consistency in grading throughout the department; the papers that were regraded eight months later showed a similar consistency for individual staff members. The exit essay experiment was felt to have been worthwhile, in part because of the cooperative effort involved in carrying it through. This led to an awareness of what the department grading standards were, of the extent to which they were being followed, and of the way in which individual grading policies compared with those of colleagues. (AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED137782

Towards Uniformity in Grading Standards

Libby Bay and Elizabeth McCulloch

Like the weather, grading is a subject everybody talks about but nobody does anything. Two years ago the English Department at Rockland Community College decided that, dire predictions to the contrary, we would seed our academic clouds and see if we could work towards some predictability in grading.

Our concern began with the depressed state of student writing and the elevated state of student grades. Sixty percent of our students earned recognition on the Dean's List; more than fifty percent of our freshmen received A's and B's in their English courses. Yet, somehow, these statistics did not jibe with our gut feelings about student accomplishments, especially in English--nor with what our eyes saw as we looked at student writing. Therefore, we made Freshman English, more particularly the grading of student themes, our special agenda for the next two years.

Our first venture was a grading workshop. The entire department came together to review five papers. When we were finished--and far apart on at least one--we argued the criteria we had used.

Obviously, five essays provided us with limited data and questionable results. Thus we decided to experiment with cooperative grading on a much larger scale by giving a uniform exit essay at the end of English 101 in January 1976 to approximately 1600 freshmen. This move was a bold one since ours is a department where freedom has always been the hallmark. We have no standard texts, no departmental tests, very little administrative supervision. We have always worked from the assumption of professional integrity and responsibility and left major decisions about class conduct to individual instructors. Thus we emphasized that this venture was only an experiment, that we had no predetermined results in mind, and that the essay would have no effect on course grades unless the teacher so chose.

A committee of three who were not teaching EN 101 that semester was selected

CS 203 271

to work out a set of grading criteria to be presented to the department and to choose the essay to which the students would respond. (We had unanimously decided the examination would take the form of an expository essay to be based on a short reading.) This Committee developed specific considerations for grading papers and, after the usual expressions of individual dissent, the following criteria were agreed upon, in rank order:

Content
Organization
Paragraph development
Sentence structure
Logic
Usage
Agreement and reference
Point of view
Transitional devices
Punctuation
Spelling

It is interesting to compare these criteria with those revealed in a study by the Educational Testing Service. In that project, fifty-three distinguished readers, including ten college English teachers, nine college social science teachers, eight college natural science teachers, ten writers and editors, nine lawyers and seven business executives graded three hundred freshmen themes. The scale they developed ranged, in rank order, from ideas (like our content first on the list), to mechanics (usage, punctuation and spelling-which we placed towards the bottom of our priorities), organization and analysis (somewhat higher on our scale), phrasing (which, oddly, is not really covered in our criteria), and "flavor" (style, individuality, interest, sincerity,--characteristics which we felt unmeasurable, but which obviously become determining factors in distinguishing between an "A" and a "B" paper).[1]

[1] Paul B. Diederich, John W. French, and Sydell T. Carlton, Factors in Judgments of Writing Ability (Princeton: ETS, 1962).

Then the committee looked through many essays, primarily from the Op Ed pages of the New York Times because these seemed timely, provocative and "properly sized." A satirical piece by Russell Baker entitled "School vs, Education" was chosen. The selection was, of course, kept secret, but during the semester we distributed essays of a similar type for students to discuss and write about.

We arranged with the registrar of the college to schedule all the 101 exams at the same time so that no student would have an unfair edge. Instructors were asked to do their grading (from A through F) that evening, making no marks on the students' papers, but recording the results on a roster sheet. When they turned in their papers to the Department Secretary the next morning, they were to pick up a "strange" set to grade. All English teachers, full and part-time, who were not teaching a section of 101 that semester, were also asked to mark at least one set of papers. Thus everybody in the department participated in the project.

Every student's paper, then, was seen by two teachers, his home instructor and a disinterested marker; two grades were recorded side by side on individual roster sheets along with the course grade and a notation of whether the teacher had averaged the exit essay into that semester grade in any way. Thus we developed a bank of information from which we hoped, with the help of our campus computer center, to draw information, primarily on grading consistency.

It took a while to gather the information, computerize the results, and examine their implications. In fact, while we were waiting, the Committee requested one other cooperative effort from the department. In September 1976, each full-time instructor was asked to grade again ten papers that he had done last January to test the element of self-consistency.

Naturally, we were interested in the findings of the computer. We realized that statistics, like the bed of Procrustes, can be adjusted to accommodate whatever degree of whopper we are attempting to project. Just luckily, in working out

our Exit Essay project, however, we were not trying to validate a pre-determined notion. Rather, we were simply exploring an idea. Whatever news the computer chose to deliver, we were willing to accept. And what it finally delivered was, of course, perfect floods of data which, in summary, gave us the basis for future discussions and decisions.

In all, 1569 students took the test. Of these,

96½% were passed by their own instructors;
94½% were passed by disinterested markers.

3½% were failed by their own instructors;
5½% were failed by disinterested markers.

40% received A's and B's from their own instructors;
29% received A's and B's from disinterested markers.

46% received C's and D's from their own instructors;
53% received C's and D's from disinterested markers.

Because of the size of the sample, these figures reveal a predicted level of statistical significance. Other researchers, notably Richard Braddock[2] have discovered a similar lack of correlation among readers of the same composition.

Also predictable was the discovered tendency on the part of the home instructor to grade higher than the disinterested marker who, naturally, had no personal interest in or knowledge of the student whose paper he was grading.

Still, for our purpose which, primarily, was to find out whether there was consistency in composition grading in our department, the experiment answered well. If almost 95% of students who were marked by disinterested markers passed the test, nearly a third with A's and B's, students were learning, and the disparity between the instructor's marks and the anonymous grader's marks was, on the whole, slight.

On the subject of disparity in grading amongst department members, we discovered that the disinterested marker graded one grade lower than the home

[2] Richard Braddock, Richard Lloyd-Jones, and Lowell Schoer, Research in Written Composition (Champaign, Ill.: NCTE, 1963).

instructor on 27% of the papers and two grades lower on 9% of the papers, whereas the home instructor graded one grade lower than the disinterested marker on 17% of the papers and two grades lower on 3%. The disinterested marker failed 5% of papers not failed by the home teachers, whereas the home teacher failed 2½% of papers <u>not</u> failed by the disinterested marker.

We felt that these differences were not really significant, allowing for, as they seemed to, the subjectivity of individual instructors and the lack of a personal factor in the grading by the disinterested markers.

Now, how did the grades the students received for the course compare with the grades they received in the Exit Essay?

95½% of students who passed the Exit Essay as graded by their home teachers were also passed for the semester. 94½% of students who passed the Exit Essay as graded by the disinterested marker also passed the course, a difference of 1%. Thus, it appeared to us that not only was there a general consistency in grading throughout the department, but that the instrument chosen was a fair measure of the diverse approaches we used in Freshman English 101.

A further measure offered some insight into our own performances as graders. As previously mentioned, in September of this year, 240 Exit Essays, written the previous January, were distributed to twenty-four teachers, ten papers each. Each of the participating teachers had previously graded the same ten papers, and were asked to re-grade them without being told what marks they had given them in January.

Of the 240 papers, 125 were graded exactly as before.
55 were marked one or two grades higher the second time around.
57 were marked one or two grades lower the second time around.
Of the 9 papers which failed on the first grading, 8 of them failed again on the second round.

Here again, a t-test, which was not done, might have shown a certain significant difference between the grading standards used by individual teachers on separate occasions. But, certainly, the disparity did not seem to indicate undue capriciousness or whimsy on the part of the teachers. In fact, the very closeness of the first and second gradings rather points to a kind of built-in consistency among individual staff members.

Most of the department members feel that the Exit Essay experiment, was worthwhile. Such studies, of course, have been done before, with somewhat similar findings, and we could have simply absorbed these. The proof, however, often lies in the doing. In the cooperative struggle, we learned something about ourselves, our values, our attitudes towards our students, our points of agreement and disagreement. We were pleased to discover that all of us, working in our different ways, are moving in the same general direction, using similar standards, attempting related goals. Despite the good feelings, the notion of an Exit Essay as standard end-of-the-semester procedure, however, was greeted with reluctance by the department.

Some members feel that making the Exit Essay a permanent, mandatory part of the Freshman English curriculum will lead to standardization of the course, to lock-step teaching for the exam, to an invasion of professionalism, and to an intrusion of privacy. Others, more practical, see the problem as one of difficulty in choosing a suitable essay. Russell Baker's essay, was conceived by several teachers to have been a poor selection. Some thought that Baker's satire took unfair advantage of the students' lack of sophistication; others that the question called for too much re-capping, a writing device which composition teachers attempt, often vainly, to train out of their students. Yet, even if a common Exit Essay is never again undertaken at Rockland

Community College, we have had a consciousness-raising experience and developed an awareness of what department standards are, whether we are following them, and how our individual grading policies compare with our colleagues.

This experiment was deliberately limited in scope and made no attempt to tackle far more important questions: are our students learning to think and write during the year of freshman composition? how different are their ideas and expression when they leave from when they came? has their humanity, in some way, been touched by their stay with us?

Our attempts to move towards uniformity in grading standards, temporarily completed, have, perhaps, cleared the way for this other two-year project!